

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Hillman, Jennifer L.  
Goli, Surya K.
- (ii) TITLE OF THE INVENTION: NOVEL HUMAN MLS3 PROTEIN
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.  
(B) STREET: 3174 Porter Drive  
(C) CITY: Palo Alto  
(D) STATE: CA  
(E) COUNTRY: USA  
(F) ZIP: 94304
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: DOS  
(D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: To Be Assigned  
(B) FILING DATE: Herewith  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER:  
(B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Billings, Lucy J.  
(B) REGISTRATION NUMBER: 36,749  
(C) REFERENCE/DOCKET NUMBER: PF-0223 US
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 415-855-0555  
(B) TELEFAX: 415-845-4166  
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 262 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: BRAITUT02

(B) CLONE: 762280

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

Met Phe Arg Phe Met Arg Asp Val Glu Pro Glu Asp Pro Met Phe Leu
 1           5           10           15
Met Asp Pro Phe Ala Ile His Arg Gln His Met Ser Arg Met Leu Ser
          20          25          30
Gly Gly Phe Gly Tyr Ser Pro Phe Leu Ser Ile Thr Asp Gly Asn Met
          35          40          45
Pro Gly Thr Arg Ala Ala Ser Arg Arg Met Gln Gln Ala Gly Ala Val
          50          55          60
Xaa Pro Phe Gly Xaa Leu Gly Met Ser Gly Gly Phe Met Asp Met Phe
65          70          75          80
Gly Met Met Asn Asp Met Xaa Gly Asn Met Glu His Met Thr Ala Gly
          85          90          95
Gly Asn Cys Gln Thr Phe Ser Ser Ser Thr Val Ile Ser Tyr Ser Asn
          100         105         110
Thr Gly Asp Gly Ala Pro Lys Val Tyr Gln Glu Thr Ser Glu Met Arg
          115         120         125
Ser Ala Pro Gly Gly Ile Arg Glu Thr Arg Arg Thr Val Arg Asp Ser
          130         135         140
Asp Ser Gly Leu Glu Gln Met Ser Ile Gly His His Ile Arg Asp Arg
145         150         155         160
Ala His Ile Leu Gln Arg Ser Arg Asn His Arg Thr Gly Asp Gln Glu
          165         170         175
Glu Arg Gln Asp Tyr Ile Asn Leu Asp Glu Ser Glu Ala Ala Ala Phe
          180         185         190
Asp Asp Glu Trp Arg Arg Glu Thr Ser Arg Phe Arg Gln Gln Arg Pro
          195         200         205
Leu Glu Phe Arg Arg Leu Glu Ser Ser Gly Ala Gly Gly Arg Arg Ala
          210         215         220
Glu Gly Pro Pro Arg Leu Ala Ile Gln Gly Pro Glu Asp Ser Leu Pro
225         230         235         240
Asp Ser Pro Ala Ala Met Thr Gly Glu Gly Pro Gly Ala Ser Ala Leu
          245         250         255
Leu Tyr Arg Leu Arg Gly
          260

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1322 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: BRAITUT02

(B) CLONE: 762280

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGGGGGCGTA CGGAGGTGGC AGCTGTGGGA GGAGGCGGCG TGAAGGCCG AGGAGCTCAA 60  
 GCGGACCA ATCCACGT TCCGGCCGC CACCCTGACC CTGCAGCGTA CCGGAAGCG 120  
 AAACGGCCG GATGGCCGC TGAGCCGAA TCGGGCACTG TGTGGAGCCC CCTGGAGCTG 180  
 AGATCAGGAT GTTCCGCTTC ATGAGGGACG TGGAGCCTGA GGATCCCATG TTCTGATGG 240  
 ATCCCTTTGC TATTCACCGT CAGCATATGA GCCGTATGTT GTCAGGTGGC TTTGGATATA 300  
 GCGCTTCCT CAGCATCACA GATGGCAACA TGCCAGGGAC CAGGGCTGCC AGCCGCCGA 360  
 TGCAGCAGGC TGGAGCTGTC TNCCTTTG GGNTGCTGGG AATGTCGGGT GGTTCATGG 420  
 ACATGTTTGG GATGATGAAT GACATGNTTG GAAACATGGA ACACATGACA GCTGGAGGCA 480  
 ATTGCCAGAC CTCTCATCT TCCACTGTCA TCTCCTACTC CAATACGGGT GATGGTCCCC 540  
 CCAAGGTCTA CCAAGAGACA TCAGAGATGC GCTCGGCACC AGGCGGGATC CGGAGACAC 600  
 GGAGACTGT TCGGATTCA GACAGTGGAC TGGAGCAGAT GTCCATTGGG CATCACATCC 660  
 GGGACAGGC TCACATCCTC CAGCGCTCCC GAAACATCG CACGGGGGAC CAGGAGGAGC 720  
 GGCAGGACTA TATCAACCTG GATGAGAGTG AGGCCGAGC GTTTGATGAC GAGTGGCGGC 780  
 GGGAGACCTC CCGATTCCGG CAGCAGCGTC CCCTGGAGTT TCGGCGGCTT GAGTCCTCAG 840  
 GGGCTGGGG ACGAAGGGCG GAGGGCCCTC CCGCCTGGC CATCCAGGA CTTGAGGACT 900  
 CCCTTCCCGA CAGTCCCGCC GCTATGACTG GTGAGGGCCC CGGGGCCTCA GCTCTCTTGT 960  
 ACAGGCTGAG AGGCTGAGAA ATCATCCCCCT GAATAACTTT TTCCTCTCGA TTCCCATCCC 1020  
 CAATTTAATA TTAAATTAAC AGGCAAGCCG GCGCCACCT CTCCCTGGGG GTCTCAGGGA 1080  
 GAACCTTTCA CGGCACCCTT TCCCTACCTT TTCCTTCTTT AATCTCCTGG TTTACCATTG 1140  
 ATGACTTCGG CTCTGCATCT ACTTACTTGA TTTTTCATTC TGCCACTTCA TCTCAAACC 1200  
 CCCTCACCTT TCCCATCCTA CTCCTGCCAT GCATTGAAGG GTCAATGCAT TTTGGGGTGA 1260  
 GNTTNGGTTT AGGGGCCCCCT TCATNCTNA GCTACCTGGG TCTTTGCCCA ACTTTTCTCA 1320  
 GA 1322

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 268 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank  
 (B) CLONE: 1066392

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Phe Arg Met Leu Asn Ser Ser Phe Glu Asp Asp Pro Phe Phe Ser  
 1 5 10 15  
 Glu Ser Ile Leu Ala His Arg Glu Asn Met Arg Gln Met Ile Arg Ser  
 20 25 30  
 Phe Ser Glu Pro Phe Gly Arg Asp Leu Leu Ser Ile Ser Asp Gly Arg  
 35 40 45  
 Gly Arg Ala His Asn Arg Arg Gly His Asn Asp Gly Glu Asp Ser Leu  
 50 55 60  
 Thr His Thr Asp Val Ser Ser Phe Gln Thr Met Asp Gln Met Val Ser  
 65 70 75 80  
 Asn Met Arg Asn Tyr Met Gln Lys Leu Glu Arg Asn Phe Gly Gln Leu  
 85 90 95  
 Ser Val Asp Pro Asn Gly His Ser Phe Cys Ser Ser Ser Val Met Thr  
 100 105 110

Tyr Ser Lys Ile Gly Asp Glu Pro Pro Lys Val Phe Gln Ala Ser Thr  
 115 120 125  
 Gln Thr Arg Arg Ala Pro Gly Gly Ile Lys Glu Thr Arg Lys Ala Met  
 130 135 140  
 Arg Asp Ser Asp Ser Gly Leu Glu Lys Met Ala Ile Gly His His Ile  
 145 150 155 160  
 His Asp Arg Ala His Val Ile Lys Lys Ser Lys Asn Lys Lys Thr Gly  
 165 170 175  
 Asp Glu Glu Val Asn Gln Glu Phe Ile Asn Met Asn Glu Ser Asp Ala  
 180 185 190  
 His Ala Phe Asp Glu Glu Trp Gln Ser Glu Val Leu Lys Tyr Lys Pro  
 195 200 205  
 Gly Arg His Asn Leu Gly Asn Thr Arg Met Arg Ser Val Gly His Glu  
 210 215 220  
 Asn Pro Gly Ser Arg Glu Leu Lys Arg Arg Glu Lys Pro Gln Gln Ser  
 225 230 235 240  
 Pro Ala Ile Glu His Gly Arg Arg Ser Asn Val Leu Gly Asp Lys Leu  
 245 250 255  
 His Ile Lys Gly Ser Ser Val Lys Ser Asn Lys Lys  
 260 265

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1116 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank  
 (B) CLONE: 1066391

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTTATGTGTT CCCGTCCGTA CTGGAGGCTA GCTCTTGTCG CGGCCGCGGC GAGTTAACAT 60  
 CGTTTTTCCA ATCTGTCCGC GGCTGCCGCC ACCCAAGACA GAGCCAGAAT GTTCAGGATG 120  
 CTGAACAGCA GTTTTGAGGA TGACCCCTTC TTCTCTGAGT CCATTCTTGC ACACCGAGAA 180  
 AATATGCGAC AGATGATAAG AAGTTTTTCT GAACCCTTTG GAAGAGACTT GCTCAGTATC 240  
 TCTGATGGTA GAGGGAGAGC TCATAATCGT AGAGGACATA ATGATGGTGA AGATTCTTTG 300  
 ACTCATACAG ATGTCAGCTC TTTCCAGACC ATGGACCAA TGGTGTCAAA TATGAGAAAC 360  
 TATATGCAGA AATTAGAAAG AAACCTCGGT CAACCTTCAG TGGATCCAAA TGGACATTCA 420  
 TTTTGTTCTT CCTCAGTTAT GACTTATTCC AAAATAGGAG ATGAACCGCC AAAGGTTTTT 480  
 CAGGCCTCAA CTCAAACCTCG TCGAGCTCCA GGAGGAATAA AGGAAACCAG GAAAGCAATG 540  
 AGAGATTCTG ACAGTGGACT AGAAAAAATG GCTATTGGTC ATCATATCCA TGACCGAGCT 600  
 CATGTCATTA AAAAGTCAA GAACAAGAAG ACTGGAGATG AAGAGGTCAA CCAGGAGTTC 660  
 ATCAATATGA ATGAAAGCGA TGCTCATGCT TTTGATGAGG AGTGGCAAAG TGAGGTTTTG 720  
 AAGTACAAAC CAGGACGACA CAATCTAGGA AACACTAGAA TGAGAAGTGT TGGCCATGAG 780  
 AATCCTGGCT CCCGAGAAC TAAAAGAAGG GAGAAACCTC AACAAAGTCC AGCCATTGAA 840  
 CATGGAAGGA GATCAAATGT TTTGGGGGAC AAACCTCCACA TCAAAGGCTC ATCTGTGAAA 900  
 AGCAACAAAA AATAAATAGC CATGCATTTG ATTGTGTTAG TTTTGATTGT TTTAACAGTT 960  
 AGTAATGGTG CTGGGTAATA AGCATAAGAC CAATCTCTTG CTGTTAAATC AGTTCTGTCC 1020  
 TTGGCAACTT TCTTCTGATA TCTGAATGTT CATGAAGGTC CTAGCTTTAT ATTGTCCCTC 1080  
 TTTTAGGAAT AAAATTTTGA TTTTCAACAA AAAAAA 1116

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 248 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 1399745

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Phe Arg Phe Met Arg Asp Val Glu Pro Glu Asp Pro Met Phe Leu  
1 5 10 15  
Met Asp Pro Phe Ala Ile His Arg Gln His Met Ser Arg Met Leu Ser  
20 25 30  
Gly Gly Phe Gly Tyr Ser Pro Phe Leu Ser Ile Thr Asp Gly Asn Met  
35 40 45  
Pro Gly Thr Arg Pro Ala Ser Arg Arg Met Gln Gln Ala Gly Ala Val  
50 55 60  
Ser Pro Phe Gly Met Leu Gly Met Ser Gly Gly Phe Met Asp Met Phe  
65 70 75 80  
Gly Met Met Asn Asp Met Ile Gly Asn Met Glu His Met Thr Ala Gly  
85 90 95  
Gly Asn Cys Gln Thr Phe Ser Ser Ser Thr Val Ile Ser Tyr Ser Asn  
100 105 110  
Thr Gly Asp Gly Ala Pro Lys Val Tyr Gln Glu Thr Ser Glu Met Arg  
115 120 125  
Ser Ala Pro Gly Gly Ile Arg Glu Thr Arg Arg Thr Val Arg Asp Ser  
130 135 140  
Asp Ser Gly Leu Glu Gln Met Ser Ile Gly His His Ile Arg Asp Arg  
145 150 155 160  
Ala His Ile Leu Gln Arg Ser Arg Asn His Arg Thr Gly Asp Gln Glu  
165 170 175  
Glu Arg Gln Asp Tyr Ile Asn Leu Asp Glu Ser Glu Ala Ala Ala Phe  
180 185 190  
Asp Asp Glu Trp Arg Arg Glu Thr Ser Arg Phe Arg Gln Gln Arg Pro  
195 200 205  
Leu Glu Phe Arg Arg Leu Glu Ser Ser Gly Ala Gly Gly Arg Arg Ala  
210 215 220  
Glu Gly Pro Pro Arg Leu Ala Ile Gln Gly Pro Glu Asp Ser Pro Ser  
225 230 235 240  
Arg Gln Ser Arg Arg Tyr Asp Trp  
245

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1502 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: GenBank

(B) CLONE: 1399744

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTCTAAAGGG	CAGCTGTGGG	AGGAGGCGGC	GTGGAAGGCC	GAGGAGCTCA	AGCCCGGACC	60
AATCCCCACG	TTCCGGGCCG	CGACCCTGAC	CCTGCAGCGT	ACCGGGAAGC	GAAACCGGCC	120
GGATGGGCCG	CTGAGCCCGA	ATCGGGCACT	GTGTGGAGCC	CCCTGGAGCT	GAGATCAGGA	180
TGTTCCGCTT	CATGAGGGAC	GTGGAGCCTG	AGGATCCCAT	GTTCTGATG	GATCCCTTTG	240
CTATTCACCG	TCAGCATATG	AGCCGTATGT	TGTCAGGTGG	CTTTGGATAT	AGCCCCTTCC	300
TCAGCATCAC	AGATGGCAAC	ATGCCAGGGA	CCAGGCCTGC	CAGCCGCCGG	ATGCAGCAGG	360
CTGGAGCTGT	CTCCCCCTTT	GGGATGCTGG	GAATGTCGGG	TGGTTTCATG	GACATGTTTG	420
GGATGATGAA	TGACATGATT	GGAAACATGG	AACACATGAC	AGCTGGAGGC	AATTGCCAGA	480
CCTTCTCATC	TTCCACTGTG	ATCTCCTACT	CCAATACGGG	TGATGGTGCC	CCCAAGGTCT	540
ACCAAGAGAC	ATCAGAGATG	CGCTCGGCAC	CAGGCGGGAT	CCGGGAGACA	CGGAGGACTG	600
TTCGGGATTG	AGACAGTGGA	CTGGAGCAGA	TGTCCATTGG	GCATCACATC	CGGGACAGGG	660
CTCACATCCT	CCAGCGCTCC	CGAAACCATC	GCACGGGGGA	CCAGGAGGAG	CGGCAGGACT	720
ATATCAACCT	GGATGAGAGT	GAGGCCGCAG	CGTTTGATGA	CGAGTGGCGG	CGGGAGACCT	780
CCCATTCCG	GCAGCAGCGT	CCCCTGGAGT	TTCGGCGGCT	TGAGTCTCTA	GGGGCTGGGG	840
GACGAAGGGC	GGAGGGGCCCT	CCCCGCCTGG	CCATCCAGGG	ACCTGAGGAC	TCCCCTTCCC	900
GACAGTCCCG	CCGCTATGAC	TGGTGAGGGC	CCCGGGCCCT	CAGCCTCTCT	TGTACAGGCT	960
GAGAGGCTGA	GAAATCATCC	CCTGAATAAC	TTTTTCCTCT	CGATTCCCAT	CCCCAATTTA	1020
ATATTAAATT	AACAGGCAAG	CCGGCCCCCA	CCTCTCCCTG	GGGGTCTCAG	GGAGAACCTT	1080
TCACGGCACC	CTTTCCTTAC	CTTTTCCTTC	TTTAATCTCC	TGGTTTACCA	TTGATGACTT	1140
CGCCTCTGCA	TCTACTGACT	TGATTTTTTCA	TTCTGCCACT	CCATCTTCAA	ACCCCTCAC	1200
CTTTCCCATC	CTACTCCTGC	CATGCATTGA	AGGGTCAATG	CATTTTGGGG	TGAGCTCTGG	1260
GTTTAGGGGC	CCCCTCCATC	CCTCAGCTAC	CCTGGATCTT	TGCCCACCTC	TTCCCTCAGAG	1320
CCCCCACTGA	GGGGCCGTAG	CCCTATCTAG	GGCTGTGGAA	GGAGCAGACT	GGTTCCTAAC	1380
TCTCTCCCTC	CTCCTGCCCA	CACACATCAA	AAGAATCTTC	CCTACACCCT	TCTCTGCCTT	1440
TATTTTTTTGA	TTTGTGCAAC	TTGTAAGTAG	GTGTTTATGG	AATAAAGGAG	AATGGAAAAA	1500
AG						1502